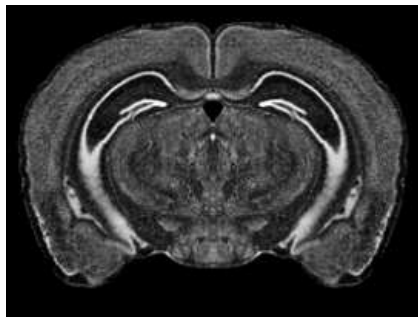


Различные способы введения текстурного потенциала для сегментации изображений.

Елшин Денис,
ВМК, ММП, 4 курс.



Биологическая мотивация:

- Статистический анализ экспрессии генов в мозге

Мотивация текстурного подхода:

- Сложность выделения границ между структурами
- Отсутствие универсальной формы анатомической структуры
- Текстурные различия соседних структур

Задача сегментации: минимизация энергии марковского поля

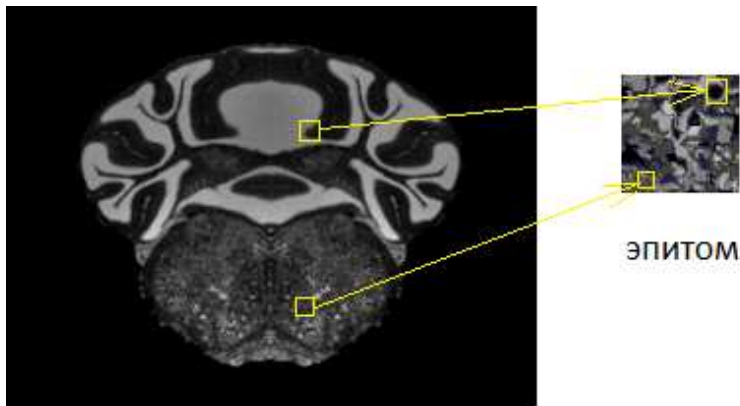
Задача:

Изображению $X = \{x_i\}_{i=1}^R$ поставить в соответствие разметку $Q = \{Q(i)\}_{i=1}^R$, $Q(i) \in \{1 \dots K\}$

Решение:

$$E(Q|X) = \sum_{i \in \mathcal{V}} \psi_i(Q(i), x_i) + \sum_{(i,j) \in \mathcal{E}} \psi_{ij}(Q(i), Q(j), x_i, x_j), \quad Q(i) = 1, \dots, K$$
$$E(Q|X) \rightarrow \min_Q$$

Эпитомный подход (Jojić, 2003)



исходное изображение

ЭПИТОМНЫЙ ПОДХОД (Joјic, 2003)

$$P(\mathcal{Z}, \mathcal{T}, \mathbf{e}) = \prod_k P(\mathbf{Z}_k, T_k, \mathbf{e}) = P(\mathbf{e}) \prod_k P(\mathbf{Z}_k | T_k, \mathbf{e}) P(T_k)$$

$$P(\mathbf{Z}_k | T_k, \mathbf{e}) = \prod_i \mathcal{N}(z_i^k | \mu_{T_k(i)}, \varphi_{T_k(i)})$$

$\mathcal{Z} = \{\mathbf{Z}_k\}_{k=1}^P$ – набор блоков изображения

\mathbf{S}_k – множество индексов пикселей блока k

T_k – преобразование пикселей k -ого блока в пиксели эпитома

$\mathcal{T} = \{T_k\}_{k=1}^P$ – набор преобразований для блоков

$\mathbf{e} = (\mu, \varphi)$ – эпитом с матожиданием μ и дисперсией φ

Эпитомный подход: сегментация изображения

$$\begin{aligned} P(\mathcal{Z}, \mathcal{Q}, \mathcal{T}, \mathbf{e}, \psi | \alpha) &= \prod_k P(\mathbf{Z}_k, \mathbf{Q}_k, T_k, \mathbf{e}, \psi | \alpha) = \\ &= P(\mathbf{e}) P(\psi | \alpha) \prod_k P(\mathbf{Z}_k, \mathbf{Q}_k | T_k, \mathbf{e}, \psi) P(T_k) \end{aligned}$$

$$P(\mathbf{Z}_k, \mathbf{Q}_k | T_k, \mathbf{e}, \psi) = \prod_i \mathcal{N}(z_i^k | \mu_{T_k(i)}, \varphi_{T_k(i)}) \psi_{T_k(i), \mathbf{Q}_k(i)}$$

$$P(\psi | \alpha) = \prod_t \text{Dir}(\psi_t | \alpha)$$

$\mathcal{Z} = \{\mathbf{Z}_k\}_{k=1}^P$ – набор блоков изображения

\mathbf{S}_k – множество индексов пикселей блока k

$\mathcal{Q} = \{\mathbf{Q}_k\}_{k=1}^P$ – набор разметок для блоков

T_k – преобразование пикселей k -ого блока в пиксели эпитома

$\mathcal{T} = \{T_k\}_{k=1}^P$ – набор преобразований для блоков

$\mathbf{e} = (\mu, \varphi)$ – эпитом с матожиданием μ и дисперсией φ

$\psi = \{\psi_{t,l}\}$ – вероятности метки класса l для элемента эпитома t

Обучение: EM-алгоритм

$$P(\mathbf{e}, \psi | \mathcal{Z}, \mathcal{Q}, \alpha) \rightarrow \max_{\mathbf{e}, \psi}$$

E-шаг:

$$q(T_k) = P(T_k | \mathcal{Z}_k, \mathcal{Q}_k, \mathbf{e}, \psi) \propto \prod \mathcal{N}(\mathbf{z}_i^k | \mu_{T_k(i)}, \varphi_{T_k(i)}) \psi_{T_k(i), \mathcal{Q}_k(i)}$$

M-шаг:

$$\mu_t = \frac{\sum_k \sum_i \sum_{T_k: T_k(i)=t} q(T_k) \mathbf{z}_i^k}{\sum_k \sum_i \sum_{T_k: T_k(i)=t} q(T_k)}$$

$$\varphi_t = \frac{\sum_k \sum_i \sum_{T_k: T_k(i)=t} q(T_k) (\mathbf{z}_i^k - \mu_t)^2}{\sum_k \sum_i \sum_{T_k: T_k(i)=t} q(T_k)}$$

$$\psi_{t,l} = \frac{\alpha_l - 1 + \sum_k \sum_{i: \mathcal{Q}_k(i)=l} \sum_{T_k: T_k(i)=t} q(T_k)}{\sum_j \alpha_j - K + \sum_k \sum_i \sum_{T_k: T_k(i)=t} q(T_k)}$$

Сегментация изображения

I - сегментируемое изображение, I_k - k -ый блок I , Q_k - его разметка.

Введём в энергию глобальный текстурный потенциал, определённый для блоков:

$$E(Q|I) = - \sum_k \log P(I_k, Q_k | e, \psi) + \sum_p \psi_p(Q(p), I) + \sum_{(p,q) \in \mathcal{E}} \psi_{pq}(Q(p), Q(q), I)$$

Для любого распределения $q(T_k)$ верна оценка:

$$\begin{aligned} - \sum_k \log P(I_k, Q_k | e, \psi) &\leq - \sum_k \sum_{p \in S_k} \sum_{T_k} q(T_k) [\log \psi_{T_k(p), Q(p)} + \\ &+ \log \mathcal{N}(I_k^p | \mu_{T_k(p)}, \varphi_{T_k(p)})] + \underbrace{\sum_k \mathbb{E}_{q(T_k)} \log q(T_k) + \sum_k \mathbb{E}_{q(T_k)} \log P(T_k)}_{const(Q)} \end{aligned}$$

$$\hat{E}(Q|I) \rightarrow \min_{Q, q(T_k)}$$

Модельные эксперименты



обучающее
изображение



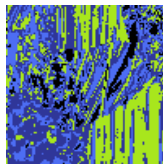
истинная
сегментация
обуч. изобр.



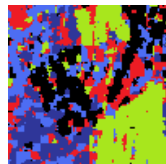
ЭПИТОМ



сегментация
эпитома



ТОЛЬКО ЦВЕТ



ТОЛЬКО ЭПИТОМЫ



тестовое
изображение



истинная
сегментация
тест. изобр.

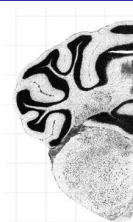


цвет+Поттс



эпитомы+Поттс

Эксперименты на срезах мозга



обучающее
изображение



истинная
сегментация



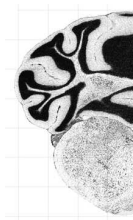
ЭПИТОМ



сегментация
эпитома



цвет+полож.



тест.изобр.



истин.сегм.

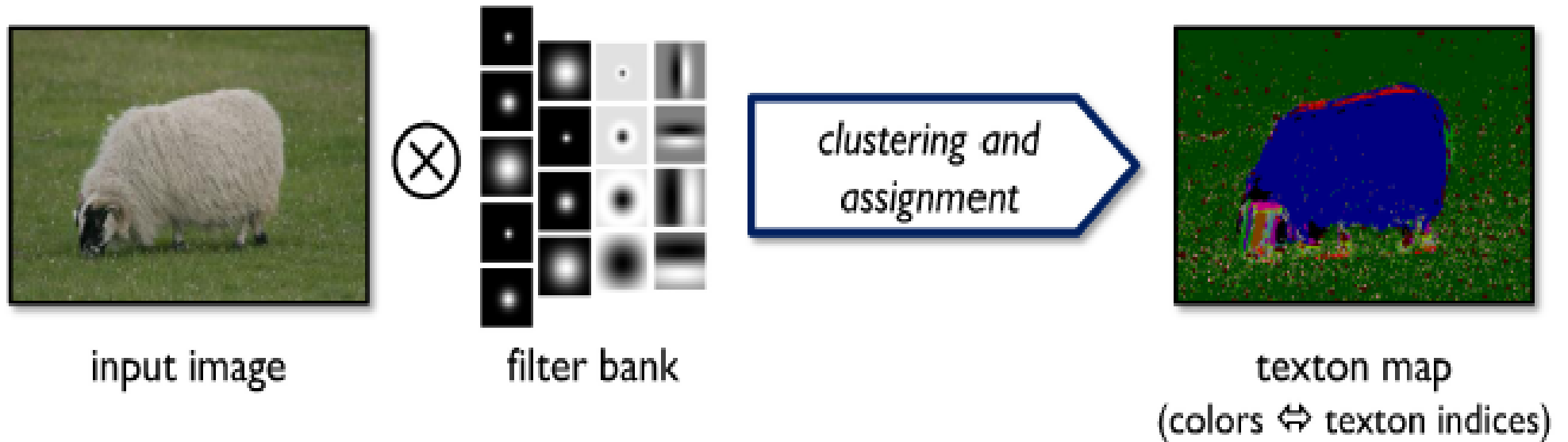


ЭПИТ.+ПОЛОЖ.



ЭПИТ.+ЦВ.+ПОЛОЖ

Textons

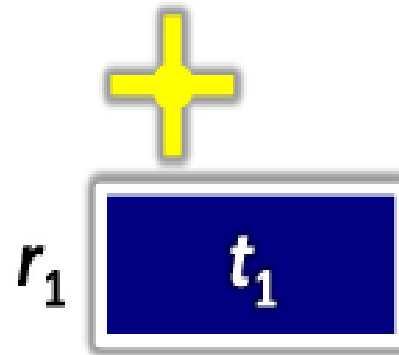


пиксель $i \rightarrow T_i \in \{1 \dots K\}$, K — число кластеров

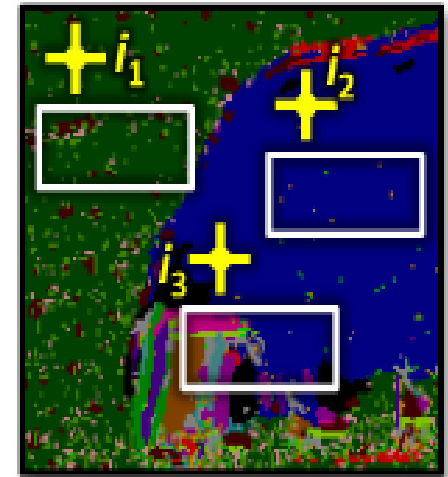
Texture-layout filters



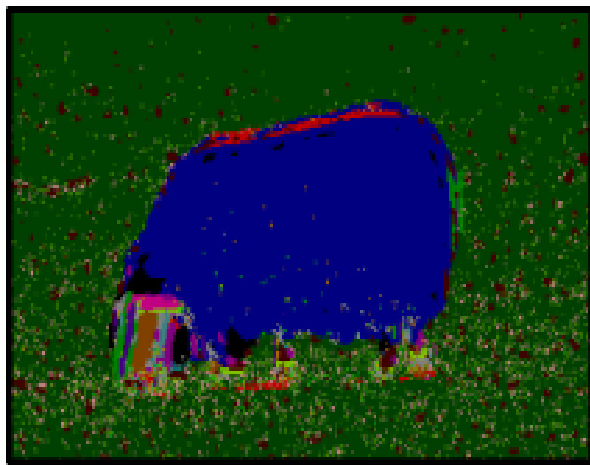
(a) input image



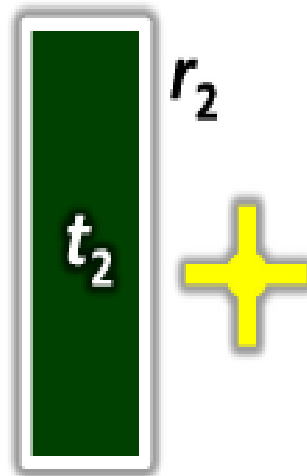
(c) feature₁ = (r_1 , t_1)



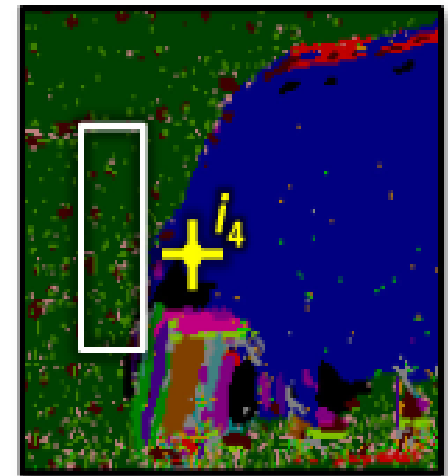
(e) feature₁ responses



(b) texture map



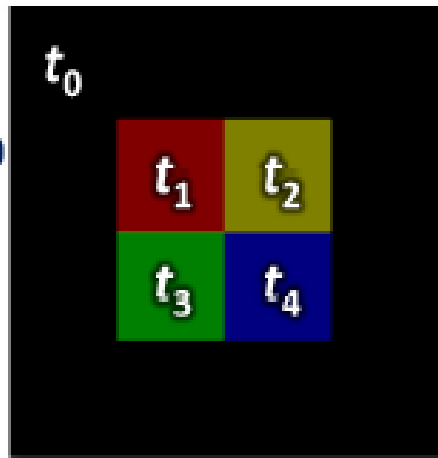
(d) feature₂ = (r_2 , t_2)



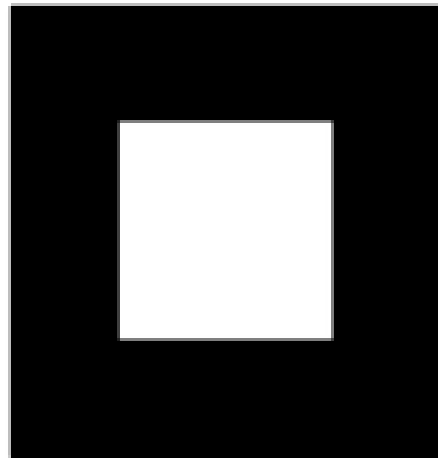
(f) feature₂ response

Texture-layout filters

training



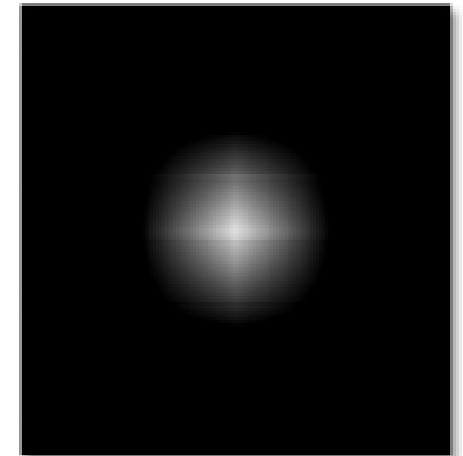
(a) training texture map



(b) ground truth labels (2 classes)

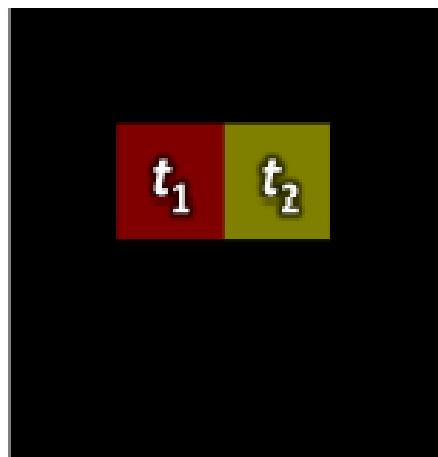


(c) texture-layout filters

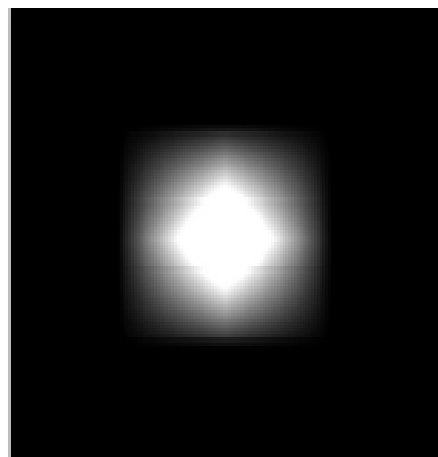


(d) feature response image
 $v_{[r_1, t_1]}(i) = v_{[r_2, t_2]}(i)$

test 1

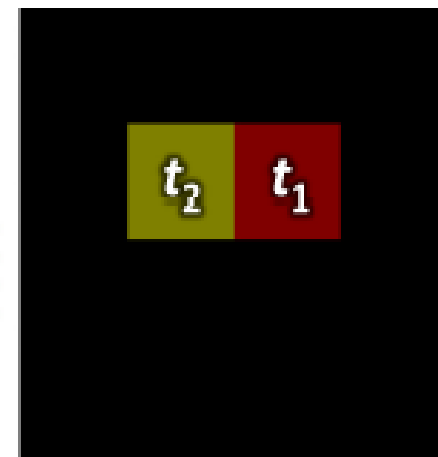


(e) test texture map

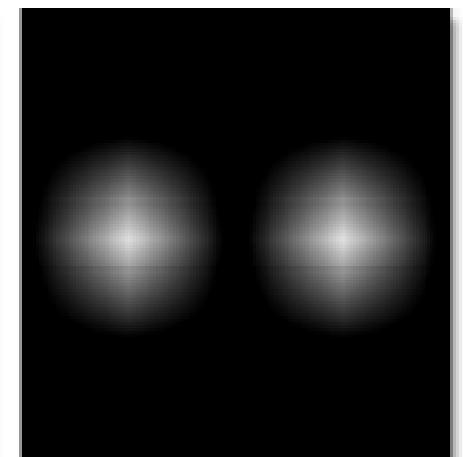


(f) $v_{[r_1, t_1]}(i) + v_{[r_2, t_2]}(i)$

test 2



(e') test texture map



(f') $v_{[r_1, t_1]}(i) + v_{[r_2, t_2]}(i)$

Texture-layout filters

(r, t) - пара (область, текстон)

$$v_{[r,t]}(i) = \frac{1}{\mathit{area}(r)} \sum_{j \in (r+i)} [T_j = t]$$

TextonBoost

$$P(c|x, i) \propto e^{H(c, i)}$$

$$H(c, i) = \sum_{m=1}^M h_i^m(c)$$

$$h_i(c) = \begin{cases} a[v_{[r, t]}(i) > \theta] + b, & c \in C \\ k^c, & \text{иначе} \end{cases}$$

Параметры бустинга: $(a, b, \{k^c\}_{c \notin C}, \theta, C, r, t)$

TextonBoost

Пиксель $i \rightarrow z_i^c \in \{-1, +1\}$, w_i^c

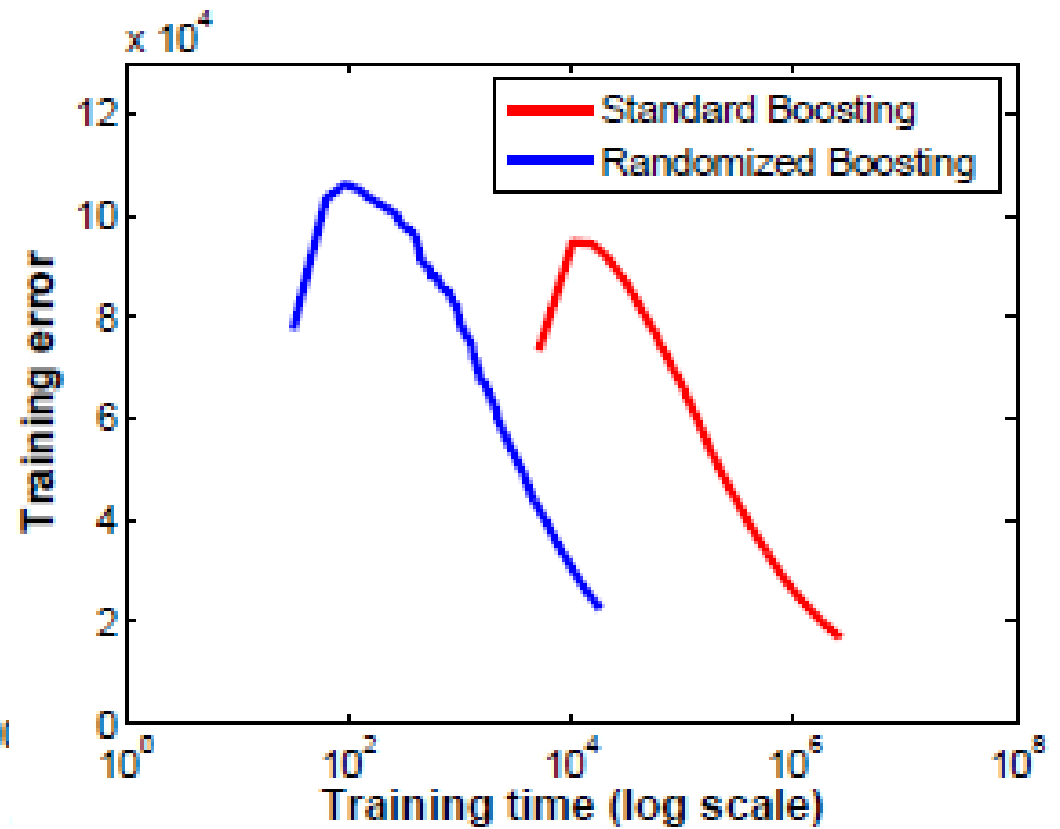
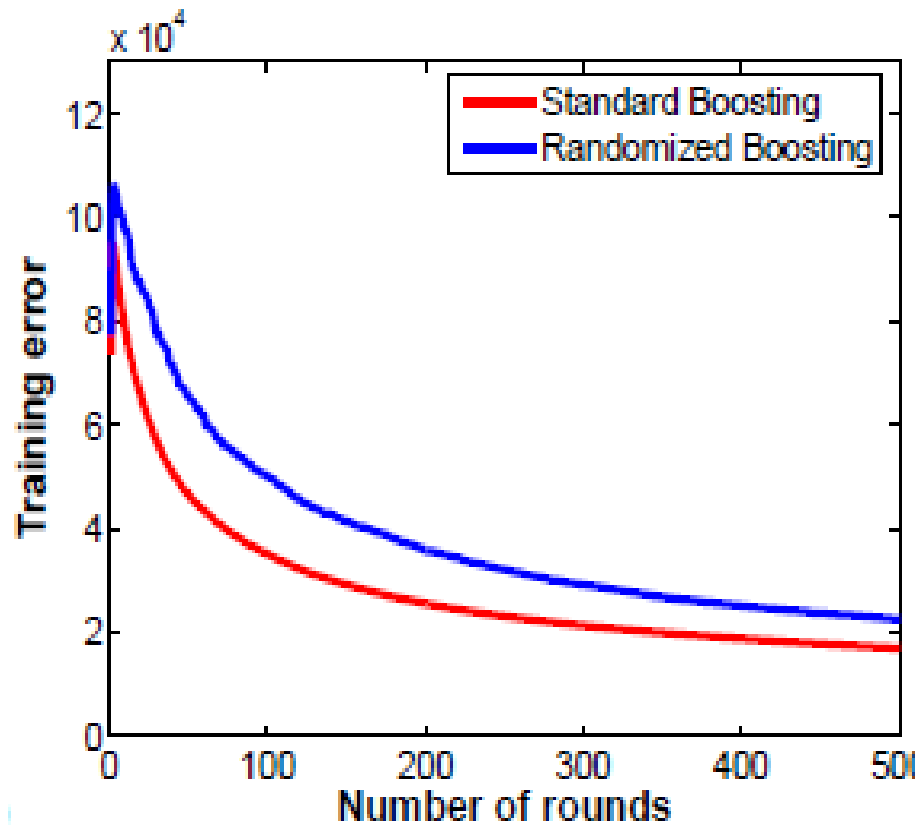
$$J_{wse} = \sum_c \sum_i w_i^c (z_i^c - h_i^m(c))^2 \rightarrow \min_{\text{boosting parameters}}$$

при известных C, r, t, θ :

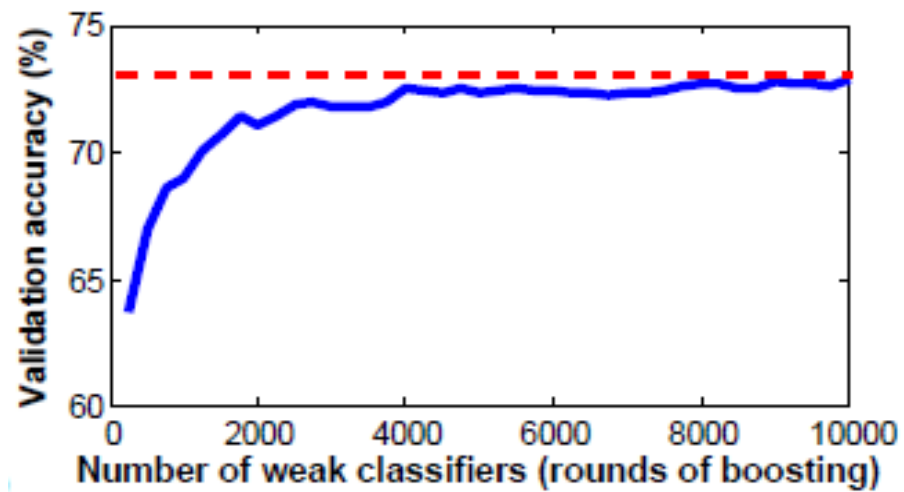
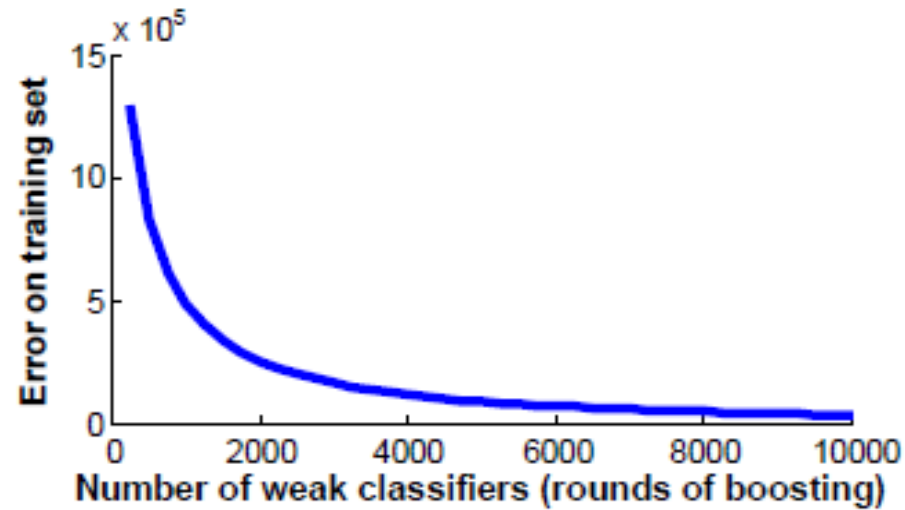
$$b = \frac{\sum_{c \in C} \sum_i w_i^c z_i^c v_{[r,t]}(i) \leq \theta}{\sum_{c \in C} \sum_i w_i^c v_{[r,t]}(i) \leq \theta}, k^c = \frac{\sum_i w_i^c z_i^c}{\sum_i w_i^c}$$

$$a + b = \frac{\sum_{c \in C} \sum_i w_i^c z_i^c v_{[r,t]}(i) > \theta}{\sum_{c \in C} \sum_i w_i^c v_{[r,t]}(i) > \theta}, w_i^c := w_i^c e^{-z_i^c h_i^m(c)}$$

TextonBoost



TextonBoost



TextonBoost



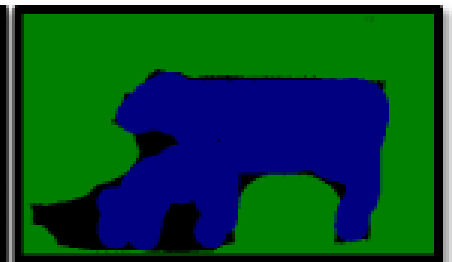
test image



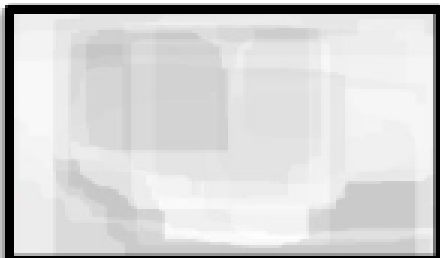
ground truth



test image



ground truth

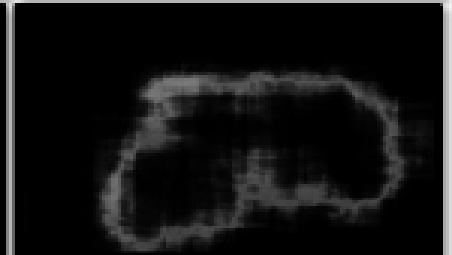
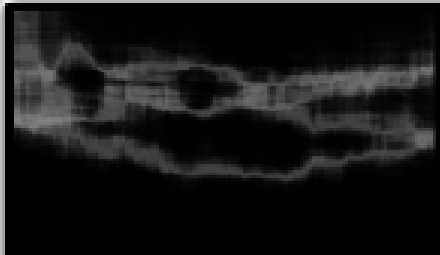
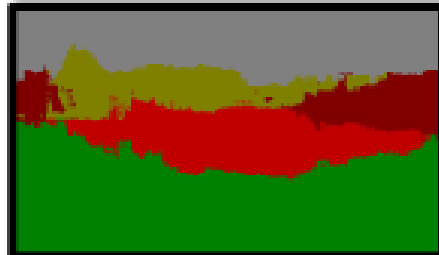
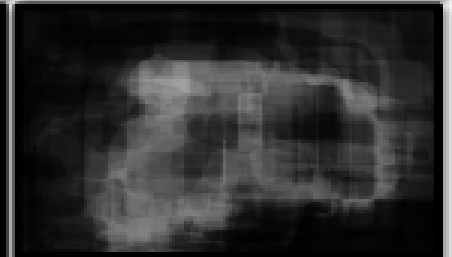
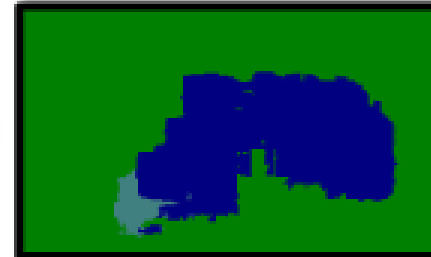
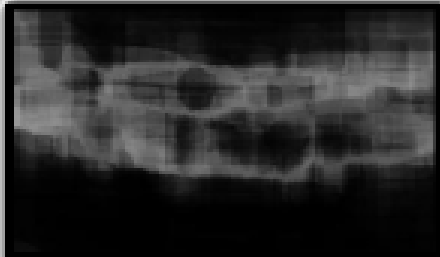
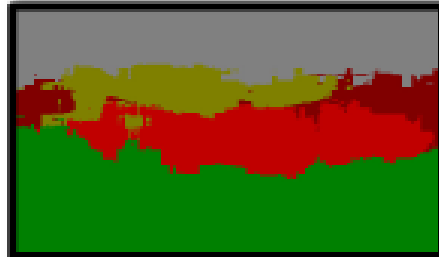
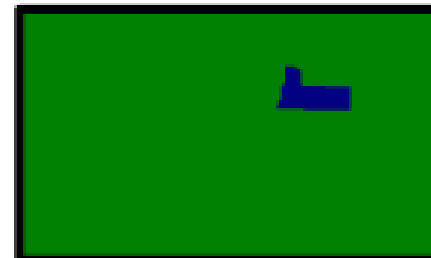


30

number of rounds

1000

2000



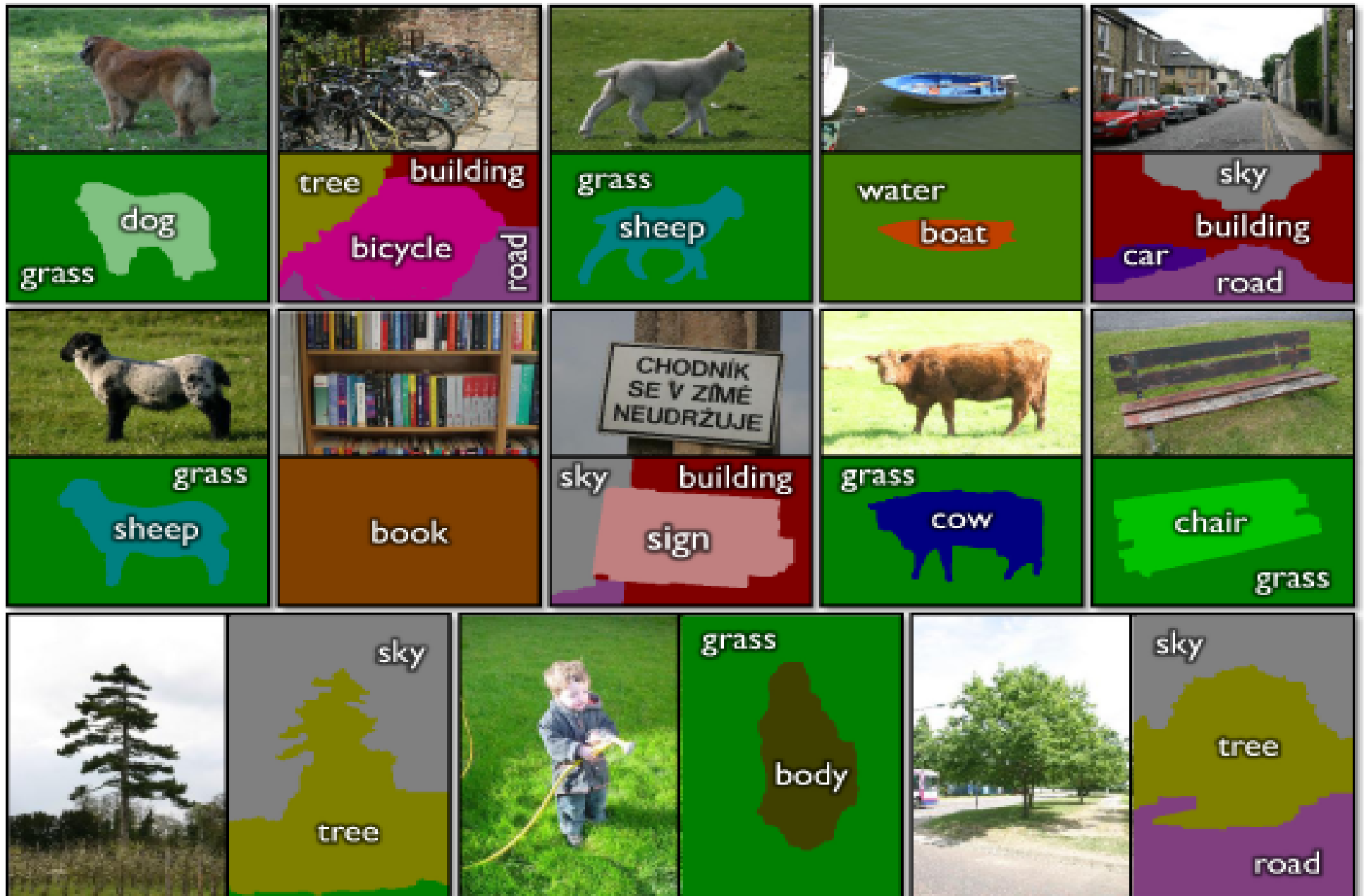
most likely label

confidence

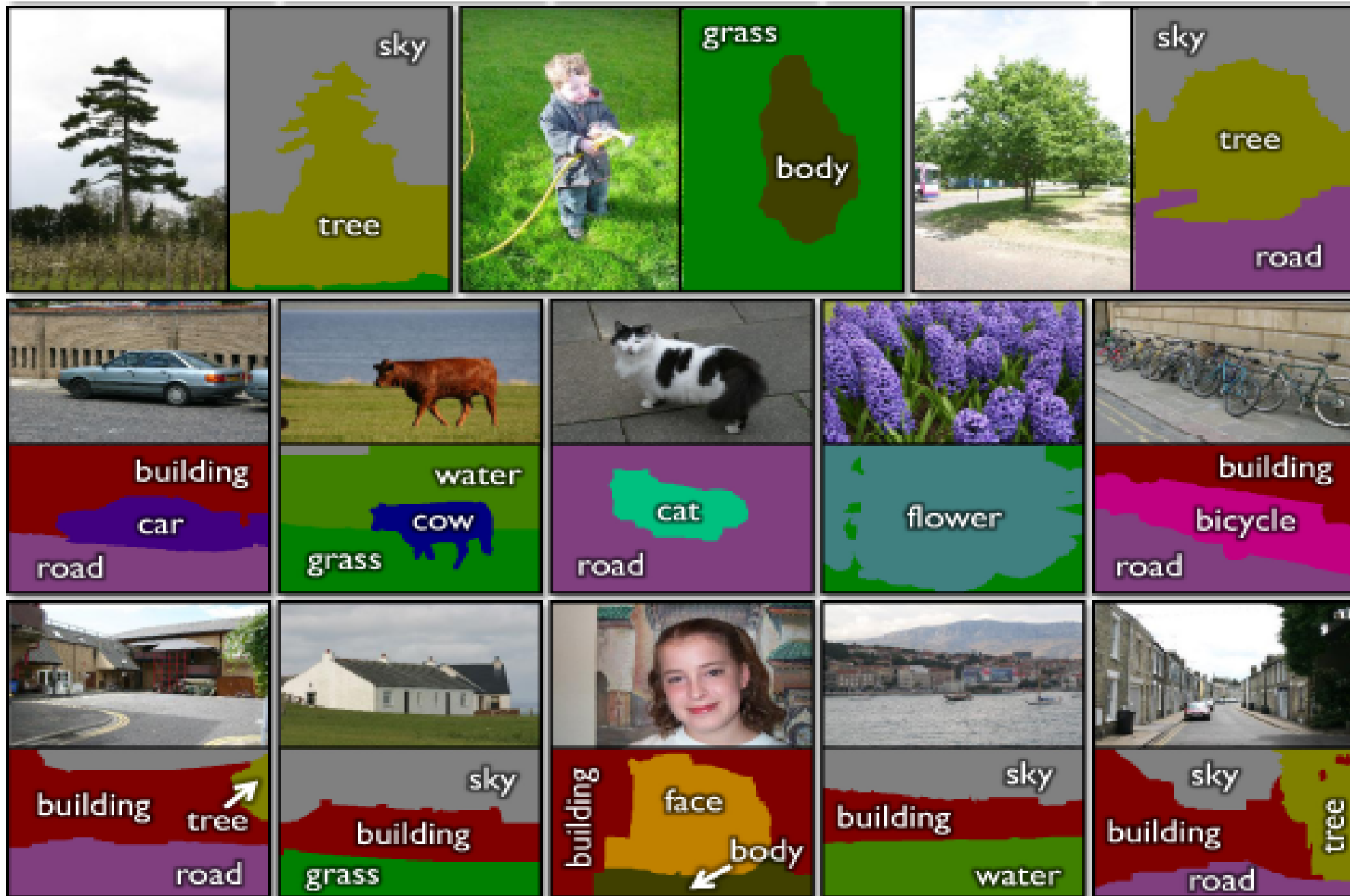
most likely label

confidence

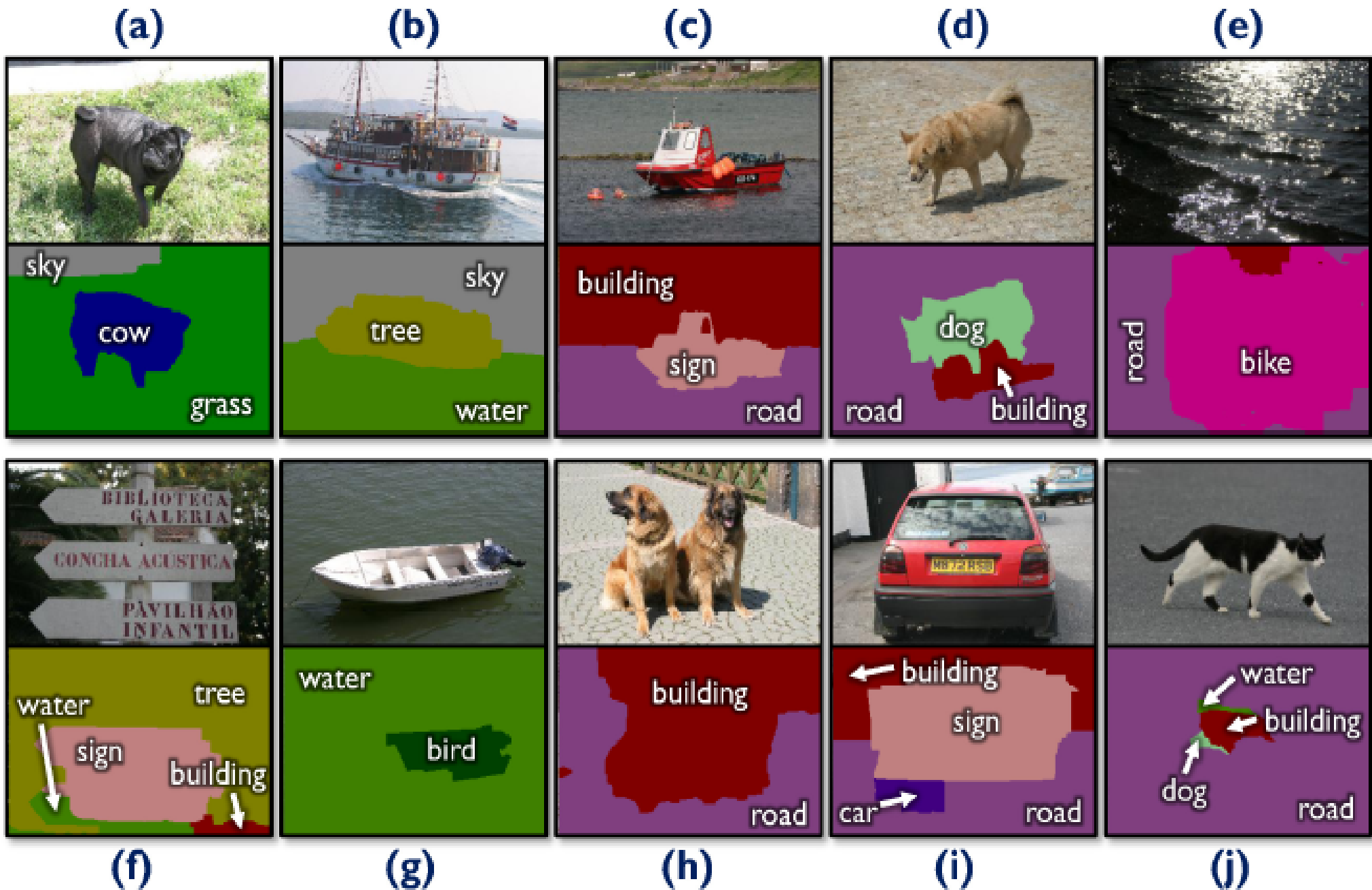
TextonBoost



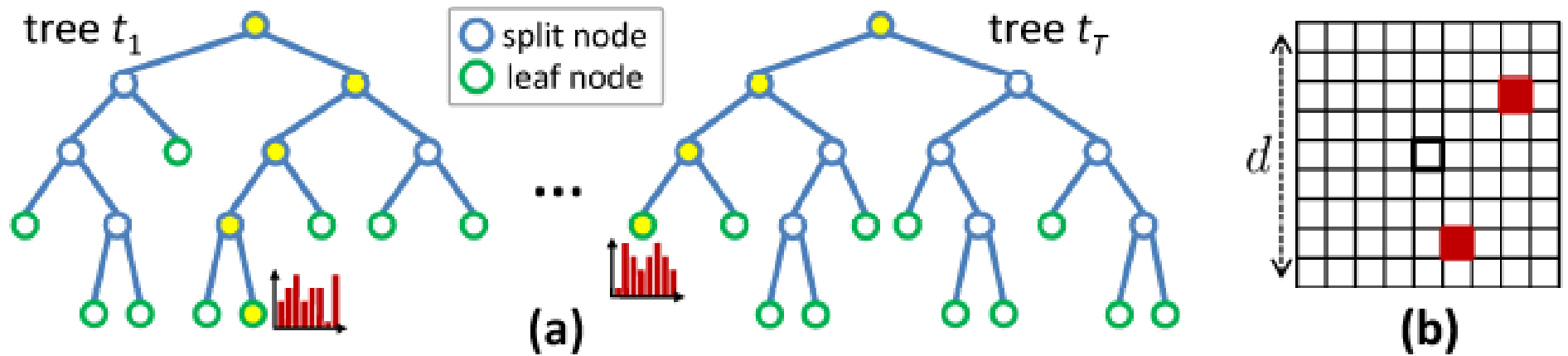
TextonBoost - эксперименты



TextonBoost - эксперименты



Решающие деревья



$$P(c|L) = \frac{1}{T} \sum_{t=1}^T P(c|l_t) \quad L = (l_1, \dots, l_T)$$

Решающие деревья: обучение

I – обучающая выборка, $I' \subseteq I$,

в вершине n обучающая выборка I_n разделяется на

$$I_l = \{i \in I, f(v_i) < t\},$$

$$I_r = I_n \setminus I_l$$

$f(v)$ выбирается из некоторого множества F

$$\text{как } \operatorname{argmax} \Delta E = -\frac{|I_l|}{|I_n|} E(I_l) - \frac{|I_r|}{|I_n|} E(I_r),$$

$E(I)$ – энтропия Шеннона

Semantic Texton Forests

$$F = \{ p_{x,y,b}; p_{x_1,y_1,b_1} \pm p_{x_2,y_2,b_2}; |p_{x_1,y_1,b_1} \pm p_{x_2,y_2,b_2}| \text{ и т.д.} \};$$

для обучения блок размера $d \times d$ пропускается
через дерево, результат:

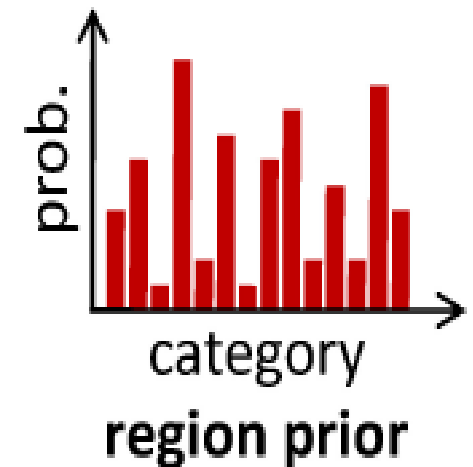
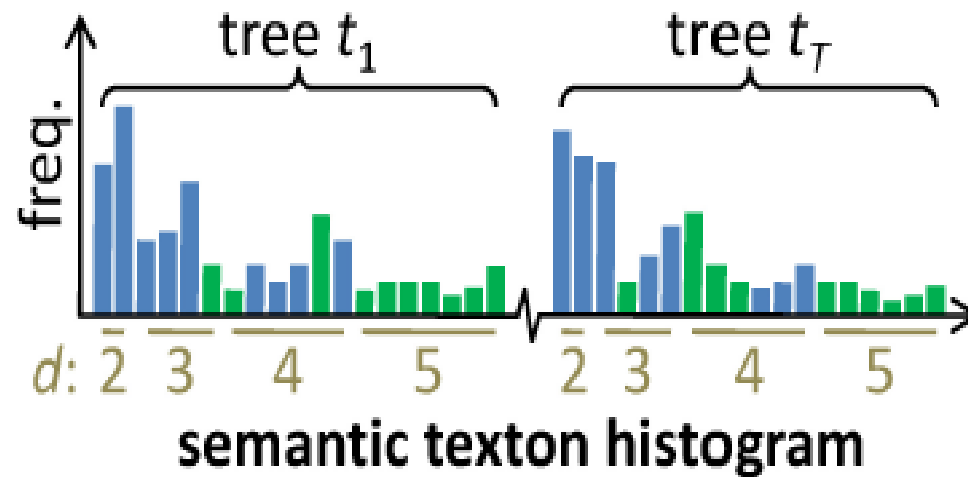
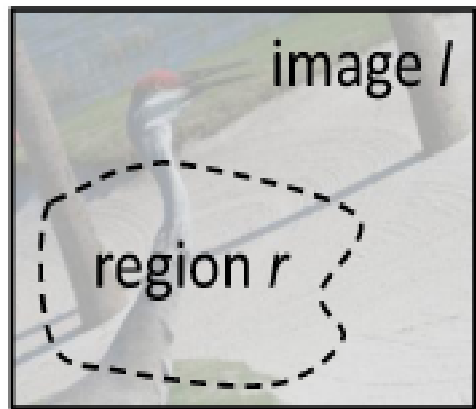
$$L = (l_1, \dots, l_T), P(c|L)$$

Semantic Texton Forests

2 вида обучения:

1. pixel-level supervision - каждому пикселю p соответствует метка класса c ;
2. weak supervision - каждому пикселю сопоставляется набор меток, присутствующих на изображении

Bag of Semantic Textons



Bags of Semantic Textons

Пиксель $i \rightarrow L_i = (l_1, \dots, l_T)_i, P(c|L_i);$

отсюда для области r извлекаем:

1. $H_r(n)$ – ненормализованную гистограмму, в которой указано, сколько раз вершина n была пройдена при пропускании r через лес,

$$H_r(n) = \sum_{n' \in \text{child}(n)} H_r(n');$$

2. Априорное распределение на классы для r :

$$P(c|r) = \sum_{i \in r} P(c|L_i)$$

Сегментация изображений

Используется два решающих леса:

1 лес - Semantic Texton forest $\rightarrow H_r, P(c|r)$;

обучение: supervised learning;

2 лес - в качестве F используем одно из семейств:

$F_1 = \{P(c=c'|r+i)\}$ – считаем долю пикселей класса c'

в прямоугольнике $r+i$;

$F_2 = \{H_{r+i}(n=n')\}$ – для прямоугольника $r+i$ считаем,

сколько пикселей дошло до вершины n' ;

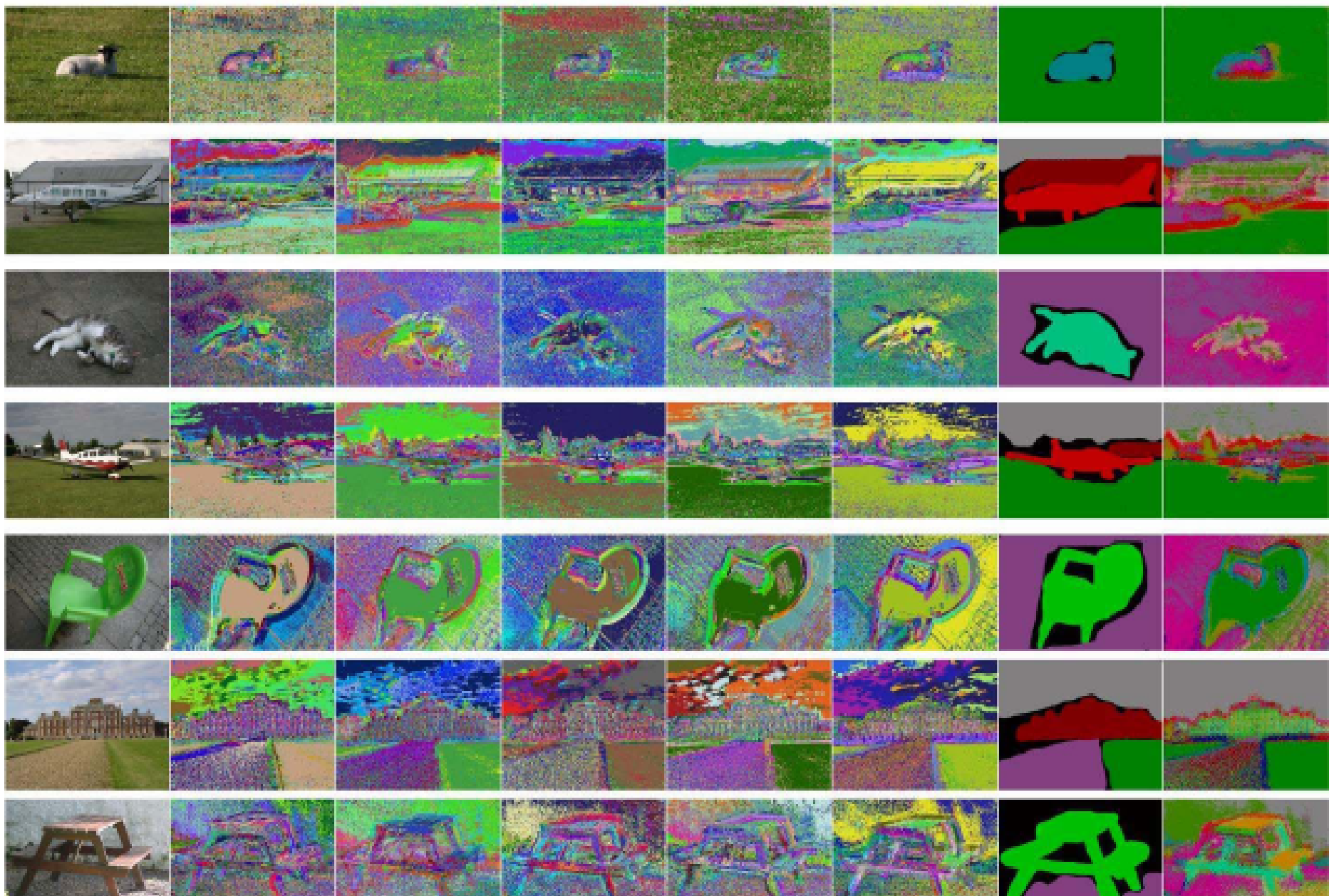
обучение - supervised learning

Сегментация изображений

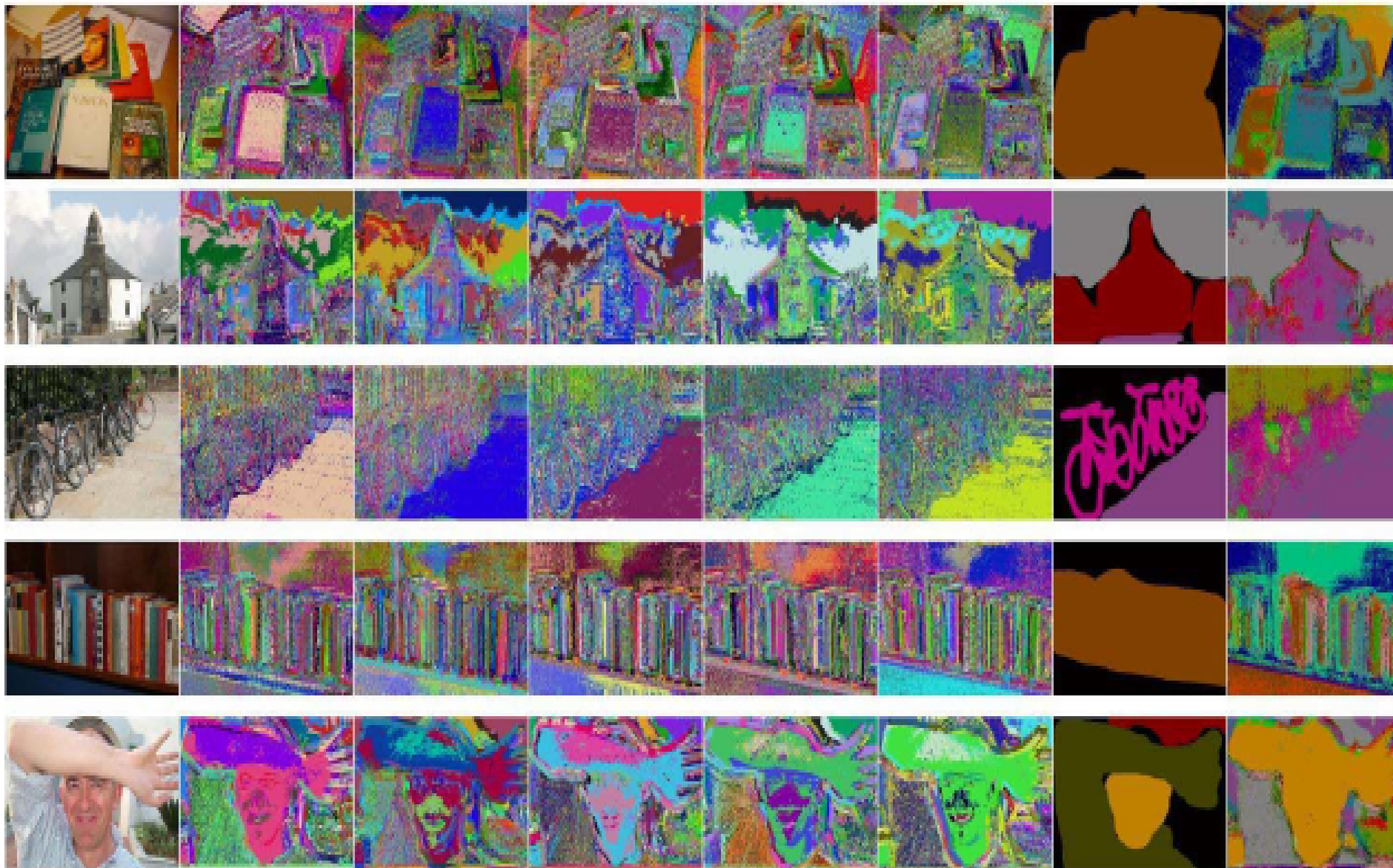
Сегментация: пропускаем изображение через 1 лес, полученные значения используем для подсчёта разделяющих функций в вершинах деревьев леса 2.

На выходе имеем $L_2 = \{l_1, \dots, l_{T_2}\}_i \rightarrow P(c|i)$

Texton Forest - эксперименты



Texton Forest - эксперименты



Категоризация изображений

Категоризация на основе SVM с ядром:

$$K(P, Q) = \frac{1}{\sqrt{Z}} \tilde{K}(P, Q)$$

$$Z = \tilde{K}(P, P) \tilde{K}(Q, Q)$$

$$\tilde{K}(P, Q) = \sum_{d=1}^D \frac{1}{2^{D-d+1}} (I_d - I_{d+1})$$

$$I_d = \sum_j \min(P_d[j], Q_d[j])$$

P, Q – гистограммы,

P_d, Q_d - гистограммы на уровне d ,

j пробегает по всем вершинам на уровне d

Категоризация изображений

$$K = \sum_t \gamma_t K_t,$$

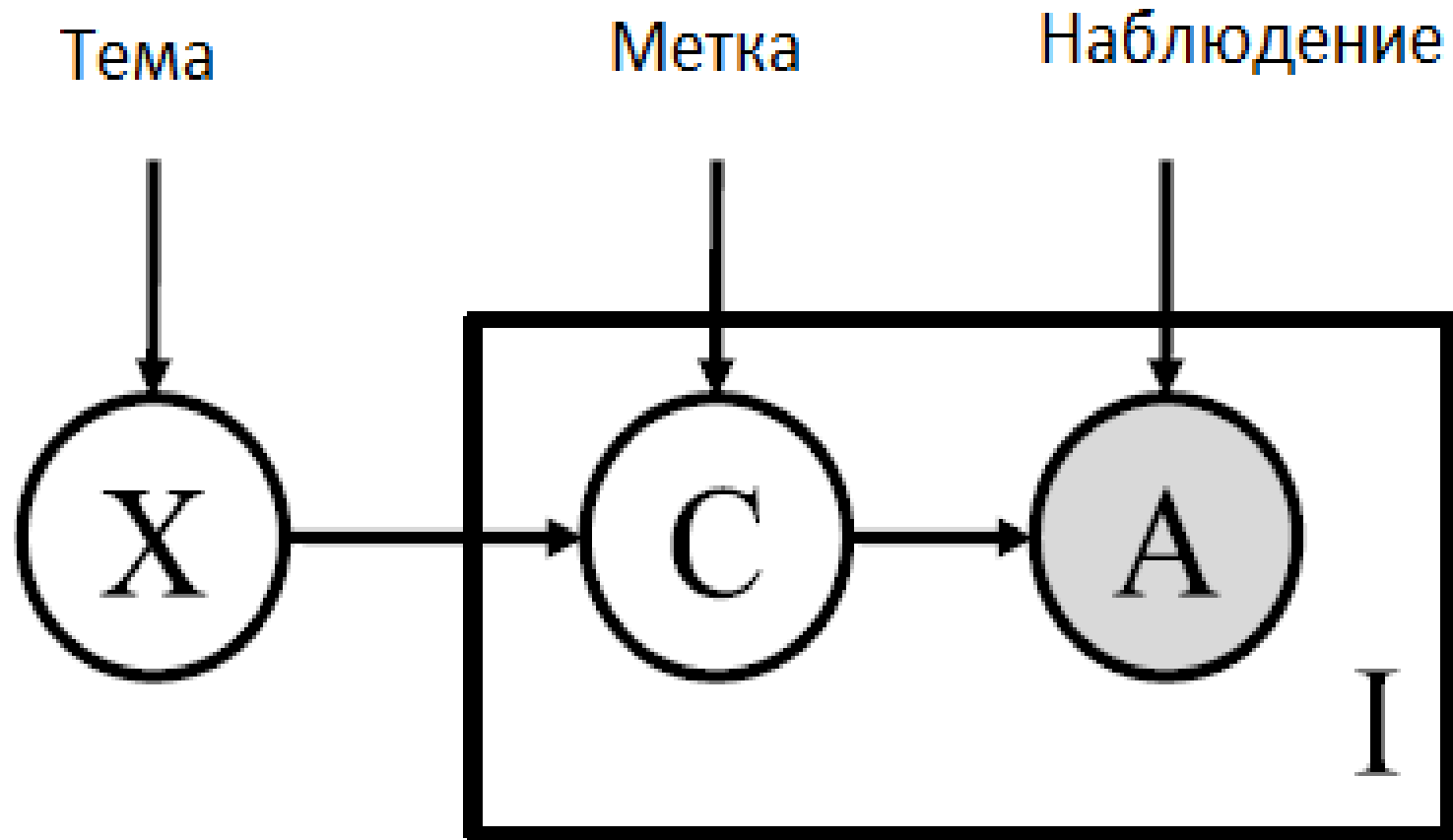
$\gamma_t = \frac{1}{T}$ — наилучший коэффициент

Категоризация изображений

1 лес, обучение - weak supervised;

затем обучаем для каждого класса SVM
с ядром K_c

Image-level prior



$$P'(c|i) = P(c|i) P(c|x)^\alpha$$

Спасибо за внимание!